THEOBROMINE SYNTHASE POLYPEPTIDE OF COFFEE PLANT AND THE GENE ENCODING SAID POLYPEPTIDE

BACKGROUND OF THE INVENTION

1. Field of the invention

[0001] This invention relates to theobromine synthase polypeptide and the gene encoding said enzyme.

2. Prior art

[0002] Coffee is a drink consumed all over the world with favorite and its utility is markedly large. On the other hand, it is known that excessive ingestion of caffeine, which is contained in coffee, causes harmful effects. Caffeine is one of xanthine derivatives and theophylline and theobromine are also the members of the xanthine derivatives. These xanthine derivatives are known to inhibit phosphodiesterase, thereby the amount of cAMP is increased. As the result, xanthine derivatives exhibit excitatory effect on the central nerves system and enhance function of the circulatory system. When they are ingested at a suitable amount, such effects of xanthine derivatives are useful for spiritual elevation. However, when the amount of digestion is excessive, they would cause harmful effects as mentioned above. Therefore, there has been a strong demand on production of a caffeine-less coffee all over the world.

[0003] To obtain caffeine-less coffee, attempts to obtain a gene involved in biosynthesis of xanthine derivatives have been performed, in the purpose to achieve artificial control of biosynthesis of caffeine. In Fig. 1 (cited from Advances in Botanical Research, Vol. 30, Academic Press (1999) p149), the pathway working for caffeine biosynthesis in coffee plants is shown. In Fig. 1, the arrow with solid line indicates the main pathway of caffeine synthesis and the arrow with dotted line indicates the minor pathway of caffeine synthesis, respectively. As shown in the second line of Fig. 1, the pathway operating for biosynthesis of caffeine from xanthosine via 7-methylxanthine and theobromine has been known, which is the main pathway for biosynthesis of caffeine biosynthesis in coffee plants. The latter half of the main biosynthesis pathway of caffeine is composed of three steps of N-methylation reactions. These N-methylation reactions have been known to be dependent on S-adenosylmethoinine. There also exists a pathway (third line in

Fig. 1) in which caffeine is biosynthesized from 7-methylxanthine via para-xanthine, but it is known that contribution of this pathway is not significant. With regard to the first methylation reaction to synthesize 7-methylxanthine, a gene encoding an enzyme responsible for said reaction has been obtained and it has been already reported (International Laid-Open Publication No. WO 97/35960). However, genes involved in the second step methylation reaction and the third step methylation reaction have not been known yet. For effective and accurate manipulation of caffeine biosynthesis, more knowledge on genes that encode enzymes involved in caffeine biosynthesis should be obtained.

SUMMARY OF THE INVENTION

[0004] The first aspect of this invention is a polypeptide consisting of an amino acid sequence defined by amino acid numbers from 1 to 378 shown in SEQ ID NO: 1 in a Sequence List. A polypeptide consisting of an amino acid sequence exhibiting at least 90% of homology with SEQ ID NO: 1 is also within the scope of this invention, so far as the polypeptide has the activity to biosynthesize theobromine using 7-methylxanthine as the substrate. Such sequence may be obtained by making deletions, insertions, substitutions or any combinations thereof in the amino acid sequence of SEQ ID NO: 1.

[0005] The second aspect of this invention is a gene consisting of a base sequence defined by base numbers from 1 to 1298 shown in SEQ ID NO: 2 in a Sequence List. A gene that hybridizes with SEQ ID NO: 2 under a stringent confdition and a gene consisting of a base sequence exhibiting at least 90% of homology with SEQ ID NO: 2 is also within the scope of this invention, so far as the gene encodes a polypeptide having the activity to biosynthesize theobromine using 7-methylxanthine as the substrate. Such sequence may be obtained by making deletions, insertions, substitutions or any combinations thereof in the base sequence of SEQ ID NO: 2.

[0006] The third aspect of this invention is a transformed plant wherein expression of said gene is inhibited in the plant to decrease biosynthesis of theobromine and a seed obtained from the transformed plant. Preferably, the plant to be transformed is selected from the group consisting of Coffea arabica, Coffea canephora, Coffea liberica and Coffea dewevrei.

[0007] The fourth aspect of this invention is a transformed plant wherein said

gene is introduced in the plant to increase biosynthesis of theobromine and a seed obtained from the transformed plant. Preferably, the plant to be transformed is selected from the group consisting of Coffea arabica, Coffea canephora, Coffea liberica and Coffea dewevrei.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] The present invention will be further explained in detail hereafter with reference to the accompanying drawings, in which:

Fig. 1 is a drawing showing the pathway of caffeine biosynthesis ;

Fig. 2 is a drawing showing base sequences of cDNAs obtained from MTL1, MTL2, MTL3 and MXMT1 ;

Fig. 3 is a drawing showing alignment of amino acid sequences obtained from MXMT1, MTL2 and MTL3 ;

Fig. 4 is a photograph showing the results of SDS-PAGE analyses performed on fusion proteins obtained from MTL2, MTL3 and MXMT1 ;

Fig. 5 is a photograph showing the results of TLC to analyze enzymatic activities of the fusion proteins obtained from MTL2, MTL3 and MXMT1 ; and

Fig. 6 is a chart showing the results of HPLC performed to identify reaction products in the enzymatic reaction mixture of the fusion protein obtained from MXMT1 identified by HPLC.

DETAILED DESCRIPTION OF THE INVENTION

[0009] The present inventors remarked an enzyme participating to the second methylation step reaction and responsible for biosynthesis of theobromine, and they have obtained the gene encoding the enzyme. The enzyme is an enzyme operating to catalyze biosynthesis of theobromine from 7-methylxanthine. Therefore, when expression of the gene encoding said enzyme in inhibited, it would result in decrease of theobromine biosynthesis. In the pathway of caffeine biosynthesis, caffeine is synthesized through N-methylation of theobromine. Then when biosynthesis of theobromine is inhibited, biosynthesis of caffeine would be inhibited as well.
As described above, theobromine and caffeine exhibit similar pharmacological effect as xanthine derivatives. Therefore, isolation of a gene encoding an enzyme, which enables concurrent manipulation of theobromine biosynthesis and caffeine biosyntheses, has a great significance. That is, if a gene encoding an enzyme responsible for the final step of caffeine biosynthesis, i.e. the third methylation step,

is isolated, then expression of the gene can be inhibited. As a result, biosynthesis of caffeine would be reduced, but biosynthesis of theobromine would not be reduced. Moreover, accumulation of theobromine is expected to occur, as the metabolism of theobromine is inhibited. Thus, considering that pharmacological effect of theobromine is similar to that of caffeine, the effect of the present invention, which relates to isolation of a gene encoding theobromine synthase, can be estimated to be significant.

[0010] The present invention relates to theobromine synthase gene derived from Coffea arabica, consisting of a base sequence defined by the base numbers 1 to 1298 shown in SEQ.ID. NO:2 in a Sequence List. As described above, in coffee plants, theobromine synthase catalyzes methylation reaction at biosynthesis of theobromine using 7-methylxanthine as the substrate. The gene defined by the base sequence described in SEQ.ID. NO:2 in a Sequence List is a gene encoding theobromine synthase having such characteristic.

[0011] According to technique of gene recombination, artificial modification can be achieved at a specific site of basic DNA, without alteration or with improvement of basic characteristic of said DNA. Concerning a gene having native sequence provided according to this invention or modified sequence different from said native sequence, it is also possible to perform artificial modification such as insertion, deletion or substitution to obtain gene of equivalent or improved characteristic compared with said native gene. Moreover, a gene with such mutation is also included in the range of this invention. That is, the gene, consisting of a base sequence hybridizes with said base sequence shown in SEQ ID NO: 2 in the sequence list under stringent condition, means a gene in which 10 or less, preferably 7 or less, and more preferably 3 or less bases of the sequence is deleted, substituted or added to the base sequence shown in SEQ ID NO: 2 in a Sequence List. Moreover, such gene exhibits homology 90% or more, preferably 95% or more and still preferably 99% or more with the base sequence shown in SEQ ID NO: 2 in a Sequence List. In addition, such gene hybridizes with the base sequence shown in the SEQ ID NO: 2 in a Sequence List under stringent condition. Such gene is also within the range of this invention so far as it encodes a polypeptide having the characteristic as theobromine synthase, that catalyzes biosynthesis of theobromine using 7-methylxanthine as the substrate.

[0012] Furthermore, this invention relates to polypeptide of theobromine synthase derived Coffea arabica, consisting of an amino acid sequence defined by the amino acid numbers from 1 to 378 shown in SEQ ID NO: 1 in a Sequence List. The polypeptide consisting of an amino acid sequence in which a part of said polypeptide defined by amino acid sequence shown in SEQ ID NO: 1 is deleted, substituted or added with another amino acid sequence means a polypeptide in which 10 or less, preferably 7 or less, and more preferably 3 or less amino acids of the sequence is deleted, substituted or added to the amino acid sequence shown in SEQ ID NO: 1 in a Sequence List. Moreover, such polypeptide exhibits homology 90% or more, preferably 95% or more and still preferably 99% or more with the amino acid sequence shown in SEQ ID NO: 1 in a Sequence List. Such polypeptide is also within the range of this invention so far as it exhibits characteristic as theobromine synthase, that catalyzes biosynthesis of theobromine using 7-methylxanthine as the substrate. Incidentally, the polypeptides shown in SEQ.ID. NO:3, SEQ.ID. NO:5 and SEQ.ID. NO:7 in a Sequence List can be obtained from coffee arabica (Coffea arabica), and the polypeptides have higher than 80% of homology compared with the amino acid sequence of SEQ.ID. NO:1 in a Sequence List. These three polypeptides did not exhibit activity as theobromine synthase, despite of high homology to SEQ.ID. NO:1 in a Sequence List.

[0013] A transformed plant, in which expression of theobromine synthetase gene described in SEQ.ID. NO:2 in a Sequence List is inhibited to decrease biosynthesis of theobromine, is also within the scope of the present invention. The theobromine synthase gene of the present invention is, as mentioned above, a gene encoding an enzyme involved in biosynthesis of theobromine in coffea arabica. Thus, by inhibiting expression of the gene according to the present invention, biosynthesis of theobromine is assumed to decrease in a plant, whereby it enables decrease of theobromine content and caffeine content in the plant. As a plant of the target in which expression of theobromine synthase gene of the present invention is inhibited, coffee plants such as Coffea arabica, Coffea canephora, Coffea liberica and Coffea dewevrei and the like can be exemplified.

[0014] In these plants, by inhibiting expression of the gene of the present invention, biosyntheses of theobromine and caffeine would be reduced. As a means for inhibiting expression of the gene of the present invention, a method utilizing an

antisense gene (antisense gene method) can be adopted. The antisense gene means a gene that expresses a base sequence complementary to mRNA, a transcription product of DNA constituting a certain gene. The transcription product of the antisense gene is complementary to an inherent mRNA, then the antisense gene can inhibit gene expression at the stage of translation. By utilizing this technique, expression of theobromine synthase gene can be inhibited.

[0015] In addition, other methods that can inhibit expression of a gene have been known. By destruction of a targeted gene, expression of the gene can be inhibited. Moreover, in a plant, technique of co-suppression (transwitch technique) has been known. According to the technique, expression of the targeted gene can be inhibited by phenomenon of gene interference, even when sense gene is introduced and over-expressed. Moreover, it has been reported in recent years that Double-stranded RNA interference (RNAi) method using a double stranded RNA is effective to inhibit expression of a gene (Chiou-Fen Chuang et al. PNAS (2000) vol. 97, 4985-4990). It has been demonstrated that a double strand RNA can inhibit expression of a gene in a sequence specific manner, according to the research mainly utilizing nematodes (C.elegans) or fruit fly. In the RNAi method, such double strand RNA is utilized and it has been recently demonstrated that the method is effective for not only nematodes or fruit fly but also for plants such as Arabidopsis thaliana Heynh. The mechanism involved in inhibition of gene expression by the RNAi method is not known yet. However, this method would enable inhibition of expression of a gene, with higher efficiency compared with the above-mentioned antisense method.

[0016] By the way, purine alkaloids such as caffeine and theobromine, can exhibit effect to avoid insects and the effect is considered to be the existence value of purine alkaloids in a plant. Thus, the gene of the present invention can be introduced in a plant and biosythesis of theobromine can be increased in the plant, whereby the plant body would exhibit insect-avoiding activity. As described above, the enzyme of the present invention is responsible for biosynthesis of theobromine using 7-methylxantine as the substrate. Therefore, it is assumed that, when the above-mentioned gene encoding the 7-methylxanthine synthase (International Laid-Open Publication WO 97/35960) and the gene of the present invention are introduced into a plant concurrently, the effect would be particularly significant. When the activity of 7-methylxanthine synthase is enhanced, the amount of substrate available

for the enzyme according to the present invention would be increased. As a result, accumulation of theobromine, which is the objective product, is expected to occur.

[0017] As a method to produce a transformant, a method generally well known in this art can be adopted. A vector available for the present invention may include plasmid vectors, for example pBI121 can be exemplified, but the scope the vector is not to be limited to them. Such vector can be introduced into, for example, Agrobacterium. Then the bacteria can be utilized for infection of callus or plantlets, resulting in production of transformed plants. Furthermore, it is possible to obtain seeds derived from such transformed plants. In Japanese Laid-Open Patent Application No. 2000-245485, the present inventors have reported a method comprising infection of an embryogenic callus of a coffee plant by Agrobacterium tumefaciens EHA101 and the method enables transformation of coffee plants with high efficacy. The method for transformation described in Japanese Laid-Open Patent Application No. 2000-245485 is assumed to be particularly useful.

EXAMPLES

(Amplification by PCR)

[0018] A pair of degenerate oligonucleotide (Forward primer, GGITGYDSIDSIGGICCIAAYAC; Reverse primer, ARIYKIYYRTRRAAISWICCIGG) was synthesized, based on the region conserved among TCS1 (Kato et al., 2000, GenBank accession no. AB031280) and two proteins (Z99708 and AC008153), with their functions unknown, of Arabidopsis thaliana. These oligonucleotides correspond to amino acid sequences of GC(A/S)(A/S)GPNT and PGSF(H/Y)(G/K)(R/N)LF, respectively. In a 25 μl of reaction mixture containing Coffea arabica cDNA and the above-mentioned primer pair, PCR was performed under the conditions described below. That is, after reaction at 94°C for one minute, 30 cycles of denaturation at 94°C for 30 seconds, annealing at 52°C for 30 seconds and extension at 72°C for one minutes was performed, which was followed by a final extension at 72°C for 7 minutes, whereby the PCR reaction was completed. The amplified cDNA fragment of about 270 base pairs was used for screening of cDNA library.

(cDNA library construction and screening)

[0019] Total RNA was extracted from young leaves of coffee (Coffea arabica) and it was purified to mRNA by oligo-dT column (Pharmacia). cDNA was

synthesized from mRNA using ZAPII cDNA synthesis kit (Stratagene), it was introduced into λZAPII vector to prepare phage library. Then cDNA library was screened using the above-mentioned amplified fragment as a probe. Thirty-five of resulting positive plaques were selected randomly and converted to plasmids, then physical mappping and partial sequencing were performed. As a result, they were clarified into 4 groups of independent clones.

[0020] Clones #1, #6, #35 and #45 were representatives of each group having the longest lengths close to full length cDNAs, and base sequences of the clones were determined. Moreover, the deduced amino acid sequences encoded by the open reading frame regions of the base sequences were determined. The base sequences determined by sequencing were shown in Fig. 2. The base sequence of cDNA obtained on the clone #45 was shown in SEQ.ID. NO:2 in a Sequence List and in Fig. 2D. The region corresponding to open reading frame of said gene ranged from base numbers 32 to 1168, and the deduced amino acid sequence encoded by said region was shown in SEQ.ID. NO:1 in a Sequence List. Moreover, the base sequence of cDNA obtained on the clone #1 was shown in SEQ.ID. NO:4 in a Sequence List and in Fig. 2A. The region corresponding to open reading frame of said gene ranged from base numbers 14 to 1171, and the deduced amino acid sequence encoded by said region was shown in SEQ.ID. NO:3 in a Sequence List. Furthermore, the base sequence of cDNA obtained on the clone #6 was shown in SEQ.ID. NO:6 in a Sequence List and in Fig. 2B. The region corresponding to open reading frame of said gene ranged from base numbers 44 to 1201, and the deduced amino acid sequence encoded by said region was shown in SEQ.ID. NO:5 in a Sequence List. Moreover, the base sequence of cDNA obtained on the clone #35 was shown in SEQ.ID. NO:8 in a Sequence List and in Fig. 2C. The region corresponding to open reading frame of said gene ranged from base numbers 45 to 1163, and the deduced amino acid sequence encoded by said region was shown in SEQ.ID. NO:7 in a Sequence List. In the following, the gene corresponds the clone #45 was designated to MXMT1, the clone #1 was designated to MTL1, the clone # 6 was designated to MTL2, and the clone #35 was designated to MTL3, respectively.

[0021] The alignment compared among amino acid sequences encoded by MXMT1, MTL1, MTL2 and MTL3 was shown in Fig. 3. As a result, it was shown that these four sequences exhibit extremely high homology. To confirm the

functions charge by these polypeptides, genes corresponding to each clone were expressed in E. coli to confirm their enzymatic activities.

(Expression of GST fused protein)

[0022] The open reading frame regions of MTL1 (Clone #1), MTL2 (Clone #6), MTL3 (Clone #35) and MXMT1 (Clone #45) were amplified by PCR (polymerase chain reaction). Then, they were optionally cloned into pGEX 4T-2 vector (Pharmacia) and E. coli (JM109) cells were transformed with the resulting plasmids. The obtained E. coli cells were cultured in LB liquid medium containing ampicillin. When OD600 of the culture reached to 0.5, IPTG (isopropyl thio-β-D-galactoside) was added to it and the final concentration of IPTG was made to 1 mM, then the mixture was further cultured at 16°C for 6 hours. E. coli was desrupted by a sonicator and the protein of the purpose was purified by glutathione Sepharose 4B as a GST (glutathione S-transferase) fusion protein. Concentration of the protein was measured by the Bradford method. Each of the GST fusion protein (500 ng) was separated by SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis), then it was stained by CBB (coumasie Brilliant Blue) to confirm purification. The purities of the resulting GST fusion proteins were analyzed by SDS-PAGE and the results were shown in Fig. 4. In Fig. 4, lane 1 shows the result of MTL2 fusion protein, lane 2 shows the result of MTL3 fusion protein, lane 3 shows the result of MXMT1 fusion protein, respectively. As a result, the resulting three fusion proteins were shown to be approximately pure.

(Measurement of enzymatic activities by thin layer chromatography)

[0023] Measurement of enzymatic activity was performed using thin layer chromatography (TLC), based on the method of Kato et al. (Plant Physiol., 1996, 98, 629-636). In concrete, the reaction mixture of 100 μl, containing 100 mM Tris-HCl (pH 7.5), 200 μM substrate (xanthine, 7-methylxanthine, theobromine, paraxanthine, theophylline), 4 μM ^{14}C-labeled S-adenosylmethionine, 200 μM $MgCl_2$, 200 ng GST fusion protein, was incubated at 27°C for 2 hours. After the reaction, the resulting mixture was extracted with 1 ml of chloroform, the chloroform layer was recovered, then chloroform was evaporated by speed back concentrator. The residue was dissolved in 5 μl of 50% methanol solution, then the solution was developed by TLC (solvent for development was water:acetic acid:n-butanol= 2:1:4, v/v/v). After the development, signal of radio activity was detected by image analyzer (Fuji BAS

2000). The result of enzymatic activity, which was measured on the fusion proteins derived from MTL2, MTL3 and MXMT1 using xanthine (X), 7-methylxanthine (7-Mx), theobromine (Tb), paraxanthine (Px) and theophylline (Tp) as the substrate, was shown in Fig. 5. From Fig. 5, it was revealed that the fusion protein derived from MXMT1 exhibited potent activity to synthesize theobromine, using 7-methylxanthine as the substrate. The fusion protein derived from MXMT1 also exhibited activity to synthesize caffine, using paraxanthine as the substrate, but its relative activity was 15% of the above-mentioned activity. On the other hand, the fusion proteins derived from MTL2 and MTL3 did not exhibit activity as a methyl transferase, using the above-mentioned compounds as the substrate.

(Enzymatic activity measurement and identification of the product by HPLC)

[0024] Using high performance liquid chromatography (HPLC), enzymatic activity of the MXMT1 fusion protein was measured and reaction product obtained from the enzymatic reaction was identified. The reaction mixture of 100 µl, containing 100 mM Tris-HCl (pH 7.5), 200 µM of substrate (7-methylxanthine, paraxanthine, theobromine), 50 µM of S-adenosylmethionine, 200 µM of $MgCl_2$, 200 ng of GST fusion protein, was incubated at 27°C for 2 hours. After incubation, the mixture was extracted with 1 ml of chloroform, the chloroform layer was recovered, then chloroform was evaporated by a speed back concentrator. The residue was dissolved in 50 µl of 12% acetonitrile. Then the solution was fractionated by HPLC (Shodex Rspak DS-613 column) provided with UV detection system. As the solution for development, 12% acetonitrile was used and the signal was detected for absorbance of 254 nm.

[0025] The result was shown in Fig. 6. The MXMT1 fusion protein was reacted with S-adenosylmethionine and 7-methylxanthine, which is the substrate and the reaction product was analyzed by HPLC. The chart exhibiting the result was shown in Fig. 6A. Moreover, theobromine was analyzed for a standard compound using HPLC and the chart exhibiting the result was shown in Fig. 6B. For preparation of negative standard, the MXMT1 fusion protein, S-adenosylmethionine and 7-methylxanthine was mixed and the reaction was immediately stopped and the chart exhibiting the result was shown in Fig. 6C. For standard products, 7-methylxantine, theobromine, paraxanthine and caffeine were analyzed by HPLC, and the chart exhibiting the result was shown in Fig. 6D. Furthermore, S-adenosylmethionine

and 7-methylxanthine was reacted with MXMT1 fusion protein and then theobromine was added to the reaction mixture. The chart exhibiting the result was shown in Fig. 6E. The peak position of the reaction product detected in Fig. 6A coincided with the position of theobromine, which was analyzed as the standard compound. In addition, when theobromine was added to the enzymatic reaction mixture, only one peak was observed. Therefore, it was shown that theobromine was formed by enzymatic reaction of the MXMT1 fusion protein, using 7-methylxantine as the substrate.

[0026] According to the present invention, the polypeptide of theobromine synthase derived from coffea arabica and the gene encoding said polypeptide were provided. As theobromine synthase participates in biosynthesis of caffeine, caffeineless coffee would be obtained by preparing a transformed plant, wherein expression of gene encoding said enzyme was inhibited.

[0027]

Sequence Listing

<110> President of Nara Institute Science and Technology

<120> Theobromine synthase polypeptide of coffee plant and the gene encoding said polypeptide

<160> 8

<210> 1

<211> 378

<212> Amino acid

<213> Caffea arabica

<400> 1

```
MELQEVLHMN EGEGDTSYAK NASYNLALAK VKPFLEQCIR ELLRANLPNI NKCIKVADLG   60
CASGPNTLLT VRDIVQSIDK VGQEEKNELE RPTIQIFLND LFQNDFNSVF KLLPSFYRKL  120
EKENGRKIGS CLISAMPGSF YGRLFPEESM HFLHSCYSVH WLSQVPSGLV IELGIGANKG  180
SIYSSKGCRP PVQKAYLDQF TKDFTTFLRI HSKELFSRGR MLLTCICKVD EFDEPNPLDL  240
LDMAINDLIV EGLLEEEKLD SFNIPFFTPS AEEVKCIVEE EGSCEILYLE TFKAHYDAAF  300
SIDDDYPVRS HEQIKAEYVA SLIRSVYEPI LASHFGEAIM PDLFHRLAKH AAKVLHMGKG  360
CYNNLIISLA KKPEKSDV                                                378
```

<210> 2

<211> 1298

<212> Nucleic acid

<213> Caffea arabica

<400> 2

```
AGCAGTCGCA ATTCGATTGT CCTGCATATG AATGGAGCTC CAAGAAGTCC TGCATATGAA   60
TGAAGGTGAA GGCGATACAA GCTACGCCAA GAATGCATCC TACAATCTGG CTCTTGCCAA  120
GGTGAAACCT TTCCTTGAAC AATGCATACG AGAATTGTTG CGGGCCAACT TGCCCAACAT  180
CAACAAGTGC ATTAAAGTTG CGGATTTGGG ATGCGCTTCT GGACCAAACA CACTTTTAAC  240
AGTGCGGGAC ATTGTGCAAA GTATTGACAA AGTTGGCCAG GAAGAGAAGA ATGAATTAGA  300
ACGTCCCACC ATTCAGATTT TTCTGAATGA TCTTTTCCAA AATGATTTCA ATTCGGTTTT  360
CAAGTTGCTG CCAAGCTTCT ACCGCAAACT CGAGAAAGAA AATGGACGCA AGATAGGATC  420
GTGCCTAATA AGCGCAATGC CTGGCTCTTT CTACGGCAGA CTCTTCCCCG AGGAGTCCAT  480
GCATTTTTTG CACTCTTGTT ACAGTGTTCA TTGGTTATCT CAGGTTCCCA GCGGTTTGGT  540
```

GATTGAATTG GGGATTGGTG CAAACAAAGG GAGTATTTAC TCTTCCAAAG GATGTCGTCC 600
GCCCGTCCAG AAGGCATATT TGGATCAATT TACGAAAGAT TTTACCACAT TTCTAAGGAT 660
TCATTCGAAA GAGTTGTTTT CACGTGGCCG AATGCTCCTT ACCTGCATTT GTAAAGTAGA 720
TGAATTCGAC GAACCGAATC CCCTAGACTT ACTTGACATG GCAATAAACG ACTTGATTGT 780
TGAGGGACTT CTGGAGGAAG AAAAATTGGA TAGTTTCAAT ATTCCATTCT TTACACCTTC 840
AGCAGAAGAA GTAAAGTGCA TAGTTGAGGA GGAAGGTTCT TGCGAAATTT TATATCTGGA 900
GACTTTTAAG GCCCATTATG ATGCTGCCTT CTCTATTGAT GATGATTACC CAGTAAGATC 960
CCATGAACAA ATTAAAGCAG AGTATGTGGC ATCATTAATT AGATCAGTTT ACGAACCCAT 1020
CCTCGCAAGT CATTTTGGAG AAGCTATTAT GCCTGACTTA TTCCACAGGC TTGCGAAGCA 1080
TGCAGCAAAG GTTCTCCACA TGGGCAAAGG CTGCTATAAT AATCTTATCA TTTCTCTCGC 1140
CAAAAAGCCA GAGAAGTCAG ACGTGTAAAA GTTTGTTTTT AGTTGGTTTT TGTGCCGTTG 1200
GGGGTCTTTC GGGTATTGTC GTTTTGTATT CGTAATAAAA GTGATGTGCA AGAATAAGAT 1260
ATTTAGTACA ATATTTTCAT AAAAAAAAAA AAAAAAAA 1298

<210> 3
<211> 385
<212> Amino acid
<213> Caffea arabica
<400> 3

MELQEVLHMN GGEGEASYAK NSSFNQLVLA KVKPVLEQCV RELLRANLPN INKCIKVADL 60
GCASGPNTLL TVWDTVQSID KVKQEMKNEL ERPTIQVFLT DLFQNDFNSV FMLLPSFYRK 120
LEKENGRKIG SCLIAAMPGS FHGRLFPEES MHFLHSSYSL QFLSQVPSGL VTELGITANK 180
RSIYSSKASP PPVQKAYLDQ FTKDFTTFLR MRSEELLSRG RMLLTCICKG DECDGPNTMD 240
LLEMAINDLV AEGRLGEEKL DSFNVPIYTA SVEEVKCMVE EEGSFEILYL QTFKLRYDAG 300
FSIDDDCQVR SHSPVYSDEH ARAAHVASLI RSVYEPILAS HFGEAIIPDI FHRFATNAAK 360
VIRLGKGFYN NLIISLAKKP EKSDI 385

<210> 4
<211> 1360
<212> Nucleic acid
<213> Caffea arabica
<400> 4

GTCCTGCATA TGAATGGAGC TCCAAGAAGT CCTGCATATG AATGGAGGCG AAGGCGAAGC 60

AAGCTACGCC AAGAATTCAT CCTTCAATCA ACTGGTTCTC GCCAAGGTGA AACCTGTCCT 120

TGAACAATGC GTACGGGAAT TGTTGCGGGC CAACTTGCCC AACATCAACA AGTGCATTAA 180

AGTTGCAGAT TTGGGATGCG CTTCCGGACC AAACACACTT TTAACCGTTT GGGACACTGT 240

ACAAAGTATT GACAAAGTTA AGCAAGAAAT GAAGAATGAA TTAGAACGTC CCACCATTCA 300

GGTTTTTCTG ACTGATCTTT TCCAAAATGA TTTCAATTCG GTTTTCATGC TGCTGCCAAG 360

CTTCTACCGC AAACTTGAGA AAGAAAATGG ACGCAAAATA GGATCGTGCC TAATAGCCGC 420

AATGCCTGGC TCTTTCCACG GCAGACTCTT CCCCGAGGAG TCCATGCATT TTTTACACTC 480

TTCTTACAGT CTTCAGTTTT TATCCCAGGT TCCCAGCGGT TTGGTGACTG AATTGGGGAT 540

CACTGCGAAC AAAAGGAGCA TTTACTCTTC CAAAGCAAGT CCTCCGCCCG TCCAGAAGGC 600

ATATTTGGAT CAATTTACGA AAGATTTTAC CACATTTTTA AGGATGCGTT CGGAAGAGTT 660

GCTTTCACGT GGCCGAATGC TCCTTACTTG CATTTGTAAA GGAGATGAAT GCGACGGCCC 720

GAATACCATG GACTTACTTG AGATGGCAAT AAACGACTTG GTTGCTGAGG GACGTCTGGG 780

GGAAGAAAAA TTGGACAGTT TCAATGTTCC AATCTATACA GCTTCAGTAG AAGAAGTAAA 840

GTGCATGGTT GAGGAGGAAG GTTCTTTTGA AATTTTATAC TTGCAGACTT TTAAGCTCCG 900

TTATGATGCT GGCTTCTCTA TTGATGATGA TTGCCAAGTA AGATCCCATT CCCCAGTATA 960

CAGCGATGAA CATGCTAGAG CAGCGCATGT GGCATCATTA ATTAGATCAG TTTACGAACC 1020

CATCCTAGCA AGTCATTTTG GAGAAGCTAT TATACCTGAC ATATTCCACA GGTTTGCGAC 1080

GAATGCAGCA AAGGTTATCC GCTTGGGCAA AGGCTTCTAT AATAATCTTA TCATTTCTCT 1140

TGCCAAAAAA CCAGAGAAGT CAGACATATA AAAGCTTGTT TTTAGTTGGT TTTTGTGTTA 1200

TGGGTTGTTT TCTGATACGG GGAAAGGATT CAGTGCGGTT GGGGTTCTAT CCGAGTATTG 1260

TACTTTTTAT ATTATTAGTT GGTGTATAAT TATTATGTTA CATTGTTATA TTCGTAATAA 1320

AAGTGACGTA CAAAAATAAA ATATTTTCAT AAAAAAAAAA 1360

<210> 5

<211> 385

<212> Amino acid

<213> Caffea arabica

<400> 5

MELQEVLHMN GGEGDASYAK NSSFNQLVLA KVKPVLEQCV GELLRANLPN INKCIKVADL 60

GCASGPNTLL TVRDIVQSID KVRQEMKNEL ERPTIQVFLT DLFQNDFNSV FMLLPSFYRK 120

LEKENGRKIG SCLIAAMPGS FHGRLFPEES MHFLHSSYSL QFLSQVPSGL VTELGITANK 180

RSIYSSKASP PPVQKAYLDQ FTKDFTTFLR IRSEELLSRG RMLLTCICKG DEFDGPNTMD 240

LLEMAINDLV VEGHLEEEKL DSFNVPIYAA SVEELKCIVE EEGSFEILYL ETFKLRYDAG 300

FSIDDDCQVR SHSPEYSDEH ARAAHVASLL RSVYEPILAN HFGEAIIPDI FHRFATNAAK 360

VIRLGKGFYN NLIISLAKKP EKSDI 385

<210> 6

<211> 1304

<212> Nucleic acid

<213> Caffea arabica

<400> 6

TTTAGCAGTC CCAATTCGAT TTATGTACAA GTCCTGCATA TGAATGGAGC TCCAAGAAGT 60

CCTGCATATG AATGGAGGCG AAGGCGATGC AAGCTACGCC AAGAATTCAT CCTTCAATCA 120

ACTGGTTCTC GCCAAGGTGA AACCTGTCCT TGAACAATGC GTAGGGGAAT TGTTGCGGGC 180

CAACTTGCCC AACATCAACA AGTGCATTAA AGTTGCGGAT TTGGGATGCG CTTCCGGACC 240

AAACACACTT TTAACAGTTC GGGACATTGT ACAAAGTATT GACAAAGTTA GGCAAGAAAT 300

GAAGAATGAA TTAGAACGTC CCACCATTCA GGTTTTTCTG ACTGATCTTT TCCAAAATGA 360

TTTCAATTCG GTTTTCATGT TGCTGCCAAG TTTCTACCGC AAACTTGAGA AAGAAAATGG 420

ACGCAAGATA GGATCGTGCC TAATAGCCGC AATGCCTGGC TCTTTCCACG GCAGACTCTT 480

CCCCGAGGAG TCAATGCATT TTTTACACTC TTCTTACAGT CTTCAATTTT TATCCCAGGT 540

TCCCAGCGGT TTGGTGACTG AATTGGGGAT CACTGCGAAC AAAAGGAGCA TTTACTCTTC 600

CAAAGCAAGT CCTCCGCCCG TCCAGAAGGC ATATTTGGAT CAATTTACGA AAGATTTTAC 660

CACATTTTTA AGGATTCGTT CGGAAGAGTT GCTTTCACGC GGCCGAATGC TCCTTACTTG 720

CATTTGCAAA GGAGATGAAT TCGACGGCCC GAATACCATG GACTTACTTG AGATGGCAAT 780

AAACGACTTG GTTGTTGAGG GACATCTGGA GGAAGAAAAA TTGGACAGTT TCAATGTTCC 840

AATCTATGCA GCTTCAGTAG AAGAATTAAA GTGCATAGTT GAGGAGGAAG GTTCTTTTGA 900

AATTTTGTAC TTGGAGACTT TTAAGCTCCG TTATGATGCT GGCTTCTCTA TTGATGATGA 960

TTGCCAAGTA AGATCCCATT CCCCAGAATA CAGCGATGAA CATGCTAGAG CAGCGCATGT 1020

GGCATCATTA CTTAGATCAG TTTACGAACC CATCCTCGCA AATCATTTTG GAGAAGCTAT 1080

TATACCTGAC ATATTCCACA GGTTTGCGAC GAATGCAGCA AAGGTTATCC GCTTGGGCAA 1140

AGGCTTCTAT AATAATCTTA TCATTTCTCT TGCCAAAAAA CCAGAGAAGT CAGACATATA 1200

AAAGCTTGTT TATAGTTGGT TTTTGTGCTA TGGTTTGTTT TCTGATACGG GGAAAGGATT 1260

TAGTGCGGTT GGGGTTCAAA AAAAAAAAAA AAAAAAAAAA AAAA 1304

<210> 7

<211> 372

<212> Amino acid

<213> Caffea arabica

<400> 7

```
MELQEVLRMN GGEGDTSYAK NSAYNQLVLA KVKPVLEQCV RELLRANLPN INKCIKVADL   60
GCASGPNTLL TVRDIVQSID KVGQEKKNEL ERPTIQIFLN DLFPNDFNSV FKLLPSFYRK  120
LEKENGRKIG SCLIGAMPGS FYSRLFPEES MHFLHSCYCL QWLSQVPSGL VTELGISTNK  180
GSIYSSKASR LPVQKAYLDQ FTKDFTTFLR IHSEELFSHG RMLLTCICKG VELDARNAID  240
LLEMAINDLV VEGHLEEEKL DSFNLPVYIP SAEEVKCIVE EEGSFEILYL ETFKVLYDAG  300
FSIDDEHIKA EYVASSVRAV YEPILASHFG EAIIPDIFHR FAKHAAKVLP LGKGFYNNLI  360
ISLAKKPEKS DV                                                      372
```

<210> 8

<211> 1316

<212> Nucleic acid

<213> Caffea arabica

<400> 8

```
CTTTGGCAGT CCCAATTTGA TTTATGTACA AGTCCTGCAT ATGAATGGAG CTCCAAGAAG   60
TCCTGCGGAT GAATGGAGGC GAAGGCGATA CAAGCTACGC CAAGAATTCA GCCTACAATC  120
AACTGGTTCT CGCCAAGGTG AAACCTGTCC TTGAACAATG CGTACGGGAA TTGTTGCGGG  180
CCAACTTGCC CAACATCAAC AAGTGCATTA AAGTTGCGGA TTTGGGATGC GCTTCTGGAC  240
CAAACACACT TTTAACAGTT CGGGACATTG TCCAAAGTAT TGACAAAGTT GGCCAGGAAA  300
AGAAGAATGA ATTAGAACGT CCCACCATTC AGATTTTTCT GAATGATCTT TTCCCAAATG  360
ATTTCAATTC GGTTTTCAAG TTGCTGCCAA GCTTCTACCG CAAACTTGAG AAAGAAAATG  420
GACGCAAAAT AGGATCGTGC CTAATAGGGG CAATGCCCGG CTCTTTCTAC AGCAGACTCT  480
TCCCCGAGGA GTCCATGCAT TTTTTACACT CTTGTTACTG TCTTCAATGG TTATCTCAGG  540
TTCCTAGCGG TTTGGTGACT GAATTGGGGA TCAGTACGAA CAAAGGGAGC ATTTACTCTT  600
CCAAAGCAAG TCGTCTGCCC GTCCAGAAGG CATATTTGGA TCAATTTACG AAAGATTTTA  660
CCACATTTCT AAGGATTCAT TCGGAAGAGT TGTTTTCACA TGGCCGAATG CTCCTTACTT  720
GCATTTGTAA AGGAGTTGAA TTAGACGCCC GGAATGCCAT AGACTTACTT GAGATGGCAA  780
TAAACGACTT GGTTGTTGAG GGACATCTGG AGGAAGAAAA ATTGGATAGT TTCAATCTTC  840
CAGTCTATAT ACCTTCAGCA GAAGAAGTAA AGTGCATAGT TGAGGAGGAA GGTTCTTTTG  900
```

AAATTTTATA CCTGGAGACT TTTAAGGTCC TTTACGATGC TGGCTTCTCT ATTGACGATG 960

AACATATTAA AGCAGAGTAT GTTGCATCTT CCGTTAGAGC AGTTTACGAA CCCATCCTCG 1020

CAAGTCATTT TGGAGAAGCT ATTATACCTG ACATATTCCA CAGGTTTGCG AAGCATGCAG 1080

CAAAGGTTCT CCCCTTGGGC AAAGGCTTCT ATAATAATCT TATCATTTCT CTCGCCAAAA 1140

AGCCAGAGAA GTCAGACGTG TAAAAGTTTG TTTTTGTGTT GGGGAAAGGA ATAAGTGCCG 1200

TTGGGGGTCT TTCGGGTATT GTGCTTTTTA TATTATATTG TTTTGTATCC GTAATAAAAG 1260

TGGTGTGTAA GAATAAGATA TTTGACATAT ATTATTTTCA AAAAAAAAAA AAAAAA 1316